Ocean Systems Engineering, *Vol. 10, No. 4 (2020) 435-449*
DOI: https://doi.org/10.12989/ose.2020.10.4.435

Sonar-based yaw estimation of target object using shape prediction on viewing angle variation with neural network

Minsung Sung[a] and Son-Cheol Yu*

Department of IT Engineering, Pohang University of Science and Technology (POSTECH), 77 Cheongam-ro, Nam-gu, Pohang, Republic of Korea

(Received September 11, 2020, Revised November 22, 2020, Accepted December 2, 2020)

Abstract. This paper proposes a method to estimate the underwater target object's yaw angle using a sonar image. A simulator modeling imaging mechanism of a sonar sensor and a generative adversarial network for style transfer generates realistic template images of the target object by predicting shapes according to the viewing angles. Then, the target object's yaw angle can be estimated by comparing the template images and a shape taken in real sonar images. We verified the proposed method by conducting water tank experiments. The proposed method was also applied to AUV in field experiments. The proposed method, which provides bearing information between underwater objects and the sonar sensor, can be applied to algorithms such as underwater localization or multi-view-based underwater object recognition.

Keywords: sonar GAN; underwater GAN; object detection; sonar simulator; underwater sonar image; underwater object detection; acoustic landmark

1. Introduction

Autonomous Underwater Vehicles (AUVs) have been widely utilized for various underwater missions (Loc *et al.* 2014, Kim *et al*. 2018b, Tang *et al*. 2019, Hong and Kim 2020). Underwater object recognition is one of the necessary algorithms for AUV operations. Research to recognize various underwater objects using information scanned by the sonar sensors has been conducted. Among sonar sensors, forward scan sonar (FSS), which provides a relatively high-resolution acoustic signal in the form of images, has been used to recognize underwater target objects such as human-made landmarks, natural terrain, divers, and sea creatures (Kim and Yu 2017, Karimanzira *et al.* 2020, Maki *et al.* 2020). Furthermore, the recognized information can be applied for AUV's navigation and obstacle avoidance and surveillance systems using AUVs (Johannsson *et al*. 2010, Kim *et al*. 2019).

One of the significant problems in recognizing a fixed underwater object with the FSS is that it is difficult to predict the direction in which the AUV will meet the target object. AUVs can estimate the distance and azimuth angle to the object using the sonar and depth sensor and can figure out the roll and pitch angles of AUV itself using the IMU or DVL. On the other hand, the yaw angle between

*Corresponding author, Associate Professor, E-mail: sncyu@postech.ac.kr
[a] Ph.D. Student, E-mail: ms.sung@postech.ac.kr

http://www.techno-press.org/?journal=ose&subpage=7 ISSN: 2093-6702 (Print), 2093-677X (Online)

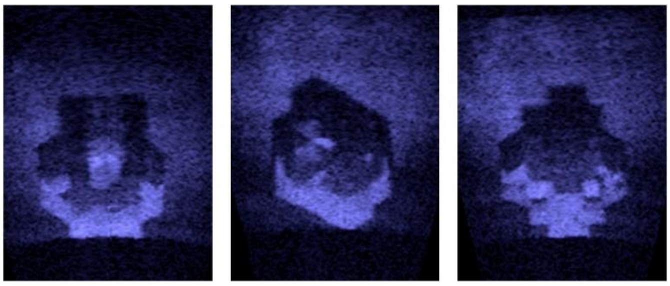


Fig. 1 Appearance of the same object in sonar images changing depending on the viewing angles

AUV and the object, which can indicate the direction in which the AUV meets the object, is challenging to measure only with on-board sensors of the AUV. However, the shape of the object changes drastically by viewing angles due to the unique imaging mechanism of the FSS, as shown in Fig. 1. Due to this characteristic of the FSS, the object recognition accuracy is limited if the yaw angle between AUV and the object is not known (Yu 2008).

If the object's yaw angle is estimated, the AUV can recognize the target object more accurately. It will also be possible to increase object recognition reliability by introducing several sonar-based algorithms that classify objects through cross-validation based on observation of the object in multiple views (Myer *et al*. 2011, Lee *et al*. 2019). Furthermore, the estimated yaw angle information can help the AUV correct its pose and position when the AUV reencounters an observed object.

This paper proposes a method to recognize the yaw angle of an underwater object using the FSS image. Furthermore, as an application of the proposed method, we verified whether the proposed method could be utilized to estimate the heading angle of AUV after deploying landmarks into the field. Aforementioned, the FSS has the characteristic that the shape of the object changes rapidly by viewing angle. This characteristic can be difficult when using the FSS for object classification. On the contrary, if all shapes of the target object according to the viewing angles can be predicted in advance, we can not only recognize the target object but also estimate the angle by comparing the shape of the object in a sonar image with the predicted shape of the object at each angle.

Addressing this feature of the FSS, the proposed method works as shown in Fig. 2. First, the sonar simulator predicts the shape according to the viewing angles of the target object. The simulated image accurately calculates the object's shape, but the real sonar images have noises due to the interference between acoustic waves and the surrounding terrains. A neural network (NN) for style transfer predicts realistic template images according to the target object's angles in sonar images by adding these effects to the simulated shapes. Finally, the proposed method calculates the similarity between the predicted template images and the underwater object's shape captured in the real sonar image. Then, by checking whether the calculated similarity exceeds a threshold and which angle of the template image has the highest similarity, we can estimate the underwater target object's yaw angle.

The proposed method has the following advantages. First, the proposed method can recognize the object's yaw angle in addition to detecting the target object. The estimated yaw angle information can be applied for the operation of AUV, such as correcting the heading angle. Next, the proposed method can estimate the yaw angle more accurately using the NN. A disadvantage of the sonar sensor is that the acoustic signal has a low signal-to-noise ratio (SNR) due to interference among acoustic beams. The proposed method employed the neural style transfer (NST) to emulate the sonar images' degradation effects and synthesize realistic template images, so the yaw angle of the underwater object can be estimated more accurately. Lastly, a template matching method considering the imaging mechanism of the sonar images was introduced. In the proposed method, the similarity between template images and real sonar images is calculated. Cross-correlation is a mainly used metric to measure the similarity between two digital images. However, the sonar image has different characteristics from the digital optical image. We calculated the similarity between two sonar images based on acoustic beams rather than pixels, focusing on the sonar sensor's imaging mechanism for more precise yaw angle estimation.

The remaining of this paper is organized as follows. Section 2 explains the proposed yaw angle estimation method through sonar image simulation, NST, and template matching in more detail. In Section 3, we describe the experiments to develop the proposed method and present the experimental results. This paper closes with the conclusion in Section 4.

2. Proposed method

2.1 Pipeline

This paper proposed a method to recognize the yaw angle of an underwater target object from sonar images. Aforementioned, the shape of an object appears differently according to the viewing angles, making it difficult to recognize the object. Taking advantage of this feature, we proposed a method to estimate the yaw angle and detect the target object by predicting the shapes of the object according to the angles in advance.

The pipeline of the proposed method to estimate the yaw angle is as follows. First, the shapes of the object according to the angles are predicted by an FSS simulator. We implemented the FSS simulator sonar by modeling the imaging mechanism of the FSS. However, the simulator approximates the imaging mechanism to calculate multiple shapes observed at various angles quickly. On the other hand, the actual sonar image has noises occurred by the interference and scattering of acoustic waves. Therefore, to estimate the angle accurately by comparing the predicted shape and the shape in the sonar image, a NN synthesizes realistic template images by applying style transfer to the predicted shapes.

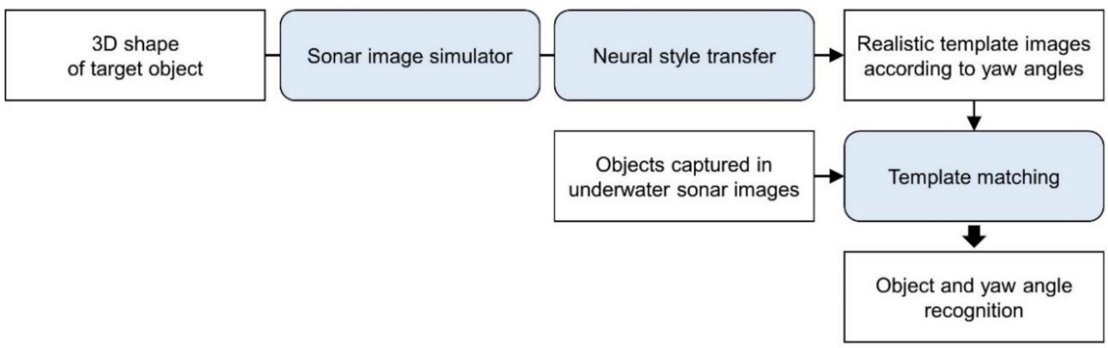


Fig. 2 The proposed method for underwater objects' yaw estimation

Finally, the underwater object's angle is estimated by calculating the similarities between the captured object in the sonar image and the template image for each angle. The proposed method can also detect the target object by checking whether the calculated similarity exceeds a threshold value.

The estimated yaw angle information can be applied to operations of AUVs. As an application, we utilized the proposed method for the localization of AUV. The angle information of the object estimated from the sonar images becomes the bearing information of the AUV based on the object. Therefore, if the AUV stores the angle of an underwater object once and reencounters the object while it operates, the AUV can correct its heading by estimating the object's angle once more. We designed several landmarks and employed them into the sea. We could then verify whether the AUV could estimate its heading from the landmark using the proposed method.

This section explains three elements of the proposed method in more details; the FSS simulator that predicts in which shape the object will be observed according to a viewing angle, the NN that synthesizes realistic template images based on the predicted shapes, and the template matching algorithm that calculates the similarity to estimate the angle finally.

2.2 FSS Simulator to Predict Shapes According to Angles

The proposed method estimates the target object's yaw angle based on the characteristic that the shapes of the object in the sonar images appear distinctively according to the viewing angle. To recognizing the yaw angle, the shape of the target object according to the viewing angles should be predicted in advance. For this purpose, we implemented the FSS simulator, which calculates the shape of given objects at the desired angle.

To implement the FSS simulator, we first analyzed the imaging mechanism of the FSS. FSS took images of an underwater scene by insonifying acoustic waves, as shown in Fig. 3. The FSS emits N acoustic beams within its azimuth range to acquire topographic information of a particular scanning area. Then, the FSS receives the echoes of the acoustic waved for a certain period. The echo of each acoustic beam creates one column of the sonar image. By mapping the intensity of the received echo to the pixel from 1 to M according to the time-of-flight (TOF) of the acoustic beam, the FSS finally obtains an M by N image for the underwater scene.

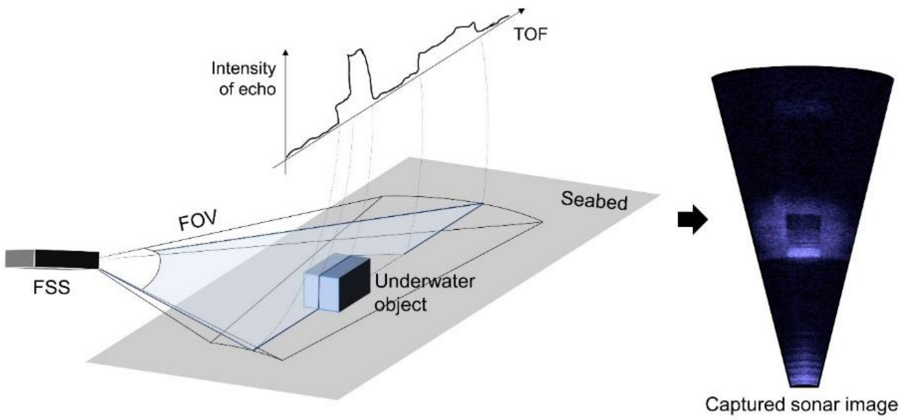


Fig. 3 Imaging mechanism of the FSS

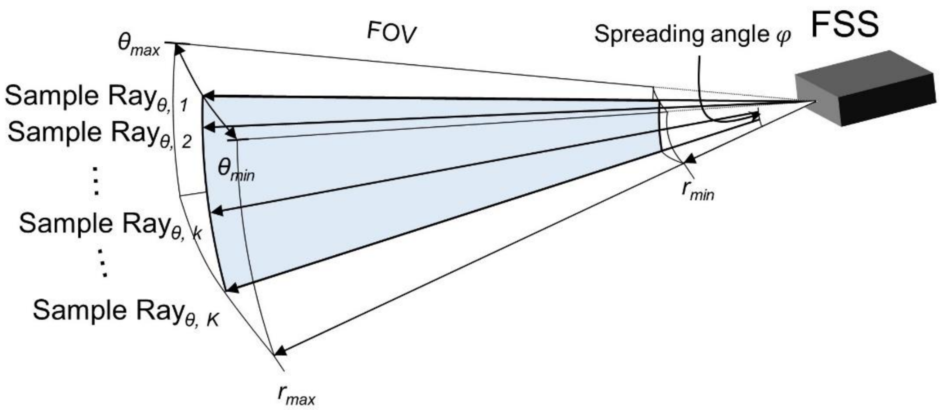


Fig. 4 Acoustic beam modeling and terminologies for the FSS simulator

The FSS simulator is implemented by emulating the imaging mechanism of the FSS (Kim *et al.* 2018a). By modeling the physical phenomena affecting an acoustic wave while the acoustic wave travels, we can calculate the shape of how a given object will appear in the sonar image. An acoustic beam transmitted from the FSS has a spreading angle φ. As shown in Fig. 4, we modeled the acoustic waves of the FSS, which has a spreading angle, using the movement of K discrete sample rays. Then, if one sample ray emitted to the azimuth angle θ is denoted as $\overrightarrow{v_{\theta,k}}$, the vector of points on the sample ray can be expressed as follows.

$$\vec{p} = t \cdot \overrightarrow{v_{\theta,k}}, \tag{1}$$

where t is a constant.

The transmitted ray is reflected off the surface of a given object and then returns to the FSS. Then, the reflection point exists on both the ray and the surface of the objects. Therefore, the reflection point $\overrightarrow{p_{\theta,k}}$ can be calculated as follows.

$$\overrightarrow{p_{\theta,k}} = \frac{\vec{N} \cdot \overrightarrow{p_1}}{\vec{N} \cdot \overrightarrow{v_{\theta,k}}} \cdot \overrightarrow{v_{\theta,k}}, \tag{2}$$

where $\vec{N}$ is a normal vector of the object surface, and $\overrightarrow{p_1}$ is a position vector of one vertex of the surface.

A sonar image is generated depending on the intensity of the echo reflected at the point $\overrightarrow{p_{\theta,k}}$ and returned to the FSS. The intensity of acoustic waves is affected by various parameters such as the material of the reflected surface and interference with other acoustic waves. However, it is difficult and makes computation complicated to model such phenomena and consider all the parameters. The proposed method requires calculating the shapes of the target object according to viewing angles, so multiple images should be simulated. To synthesize multiple images in a short time, we calculated the intensity of the echo $I_{p_{\theta,k}}$ as follows, considering only the initial intensity of the transmitted acoustic waves I_0, the transmission loss according to the distance-of-flight, and the angle of incidence of the ray α (Etter 1995).

$$I_{p_{\theta,k}} = w \frac{I_0}{\left|\overrightarrow{p_{\theta,k}}\right|^2} \cos^2 \alpha, \tag{3}$$

where w is a constant for normalization.

The intensity calculated by Eq. (3) is the intensity returned by only one sample ray. All the intensities by K sample rays should be summed up and mapped to pixel coordinates to synthesize an entire sonar image. Therefore, the sonar image I_s is constructed by the equation below.

$$I_s(r,\theta) = \sum_{1 \le k \le K, \left|\overrightarrow{p_{\theta,k}}\right| = r} I_{p_{\theta,k}}, \tag{4}$$

for $r_{min} \le r \le r_{max}$ and $\theta_{min} \le \theta \le \theta_{max}$, and where r_{min}, r_{max}, θ_{min} and θ_{max} are the parameters to define the field of view (FOV) of the FSS.

2.3 Style transfer using neural network

The shape of the given object predicted by the implemented FSS simulator is different from the real sonar images, as shown in Fig. 5. The FSS simulator considered only some significant factors and assumed the ideal situation for a quick calculation. However, in the real world, acoustic waves are affected by more factors such as interference between acoustic waves and scattering by floating materials while it travels. As a result, the actual sonar images have much noise. Although the FSS simulator accurately calculates a given object's overall shape, the output image is not realistic.

The proposed method recognizes an underwater object's yaw angle by comparing shapes in real sonar images and predicted shapes of the object according to angles. Therefore, it is necessary to make the output image of the FSS simulator realistic like the actual sonar images for more accurate angle estimation.

To make the realistic predicted image, we employed the NST. The realistic template images of the given images can also be simulated by calculating all the factors that affect acoustic waves' intensity. However, the acoustic wave's traverse is challenging to model due to several variables and phenomena (Palmese and Trucco 2006). Using NN, we could synthesize realistic template images by transferring the style of the approximately simulated images based on the distribution extracted from a large number of data without modeling.

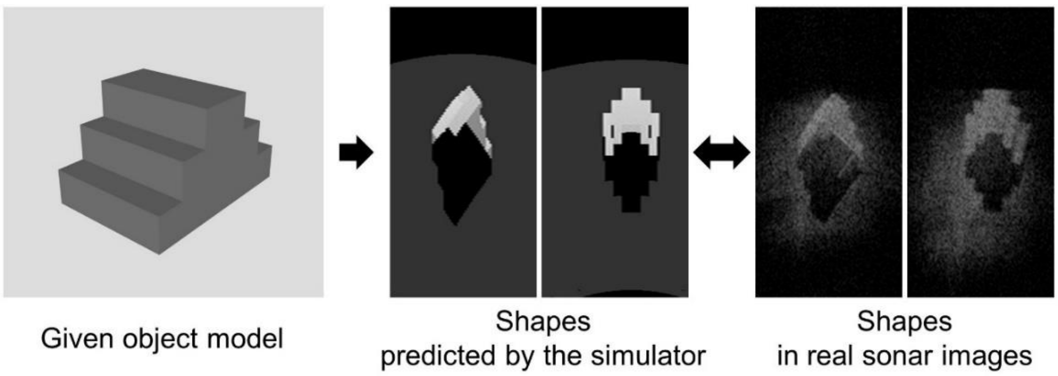
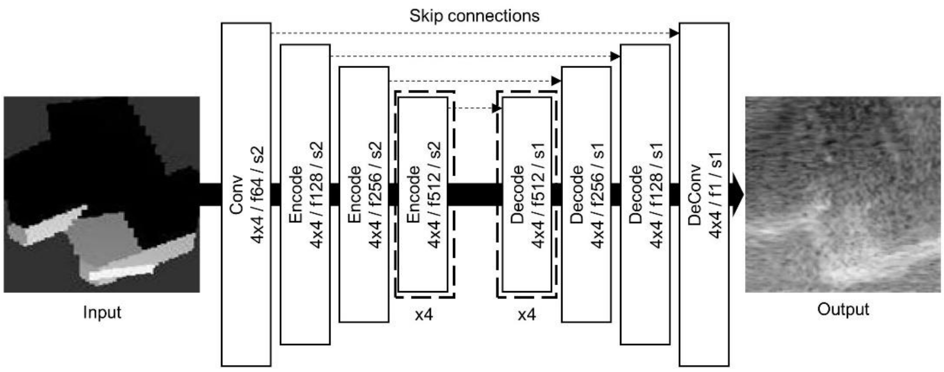
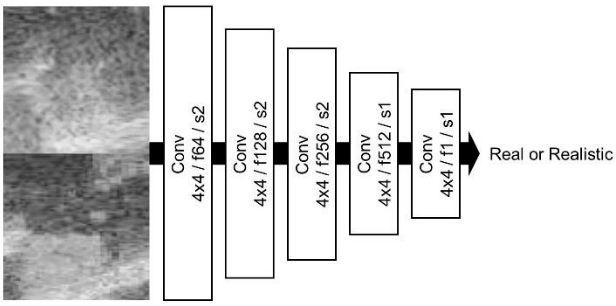


Fig. 5 Difference between the simulated shapes and real sonar images of an object



(a) Architecture of the generator



(b) Architecture of the discriminator

Fig. 6 Architecture of the neural network for simulating the realistic sonar images. 4x4 is the size of the convolutional kernel, f denotes the kernel depth, and s means the stride of the kernel

Among NN for the style transfer, we adopted the pix2pix model (Isola *et al.* 2017). Pix2pix is a generative adversarial network (GAN) to transfer an input image domain into a target domain. Using the pix2pix model, we could synthesize realistic template images of the target object by transferring the predicted shape into the domain of real sonar images.

The pix2pix network has a suitable architecture for handling sonar images. Fig. 6 describes the structure of the NN. The generator of the pix2pix model is a U-Net (Ronneberger *et al.* 2015) consisting of 16 layers. The U-Net is an encoder-decoder that has skip connections. It extracts features from the input images through the encoder. Then, the U-Net synthesized images with transferred style maintaining the contents of the input image by adding characteristics of the target domain to the extracted features through the decoder. Using the skip connections, the U-Net can transfer styles of a given image localizing the position of features more accurately.

The discriminator of the pix2pix model is a PatchGAN composed of five convolutional layers. The discriminator should determine whether the generator transferred the style of the given image realistically. Therefore, the convolutional layer, which shows outstanding performance in regression and classification problems, is used for the discriminator. Moreover, the discriminator classifies the input images by dividing the image into multiple local patches. Therefore, the generator can represent the details of the image well when transferring the style.

Finally, we designed a loss function of the pix2pix model to make the NN synthesize realistic images better. A sonar image is a form in which noise is added to the intensity of the echo reflected at the point corresponding to the specific range and azimuth angle. To make the NN predict this additive noise, we defined the loss function as follows.

$$Loss(G,D) = \mathbb{E}_{x,y}[\log D(x,y)] + \mathbb{E}_{x,z}\left[1 - \log D\big(x, G(x,z)\big)\right]$$
$$+\lambda\mathbb{E}_{x,y,z}[\|\mathcal{N}(y-x) - \mathcal{N}(G(x,z)-x)\|_1] \quad (5)$$

where G and D are the generator and discriminator respectively, x is the given simulated image, y is real sonar image corresponding to x, z is a random noise vector, and $\mathcal{N}(x)$ is a function that maps x to [-1, 1]. The first and second terms are standard loss function of the GAN. The last term gives feedback to the NN so that the effect added to the simulated image, which can also be considered as semantic information of the underwater scene, becomes similar to the real sonar image. This term allows the NN to predict the additive noise to make the simulated image look realistic.

2.4 Yaw angle estimation by template matching

The underwater target object's yaw angle can be recognized by comparing the shape of the object captured in the sea with the template images of the target object for each angle realistically synthesized through the NN. We measure the similarity between the two images to compare the object in real sonar images with the template images.

We employed a metric to calculate the similarity between two sonar images considering the characteristics of sonar images (Cho *et al.* 2015). Because one acoustic beam constitutes a column of the image according to the sonar imaging mechanism, it is practical to analyze the image in column units rather than the pixel unit. The columns of the sonar image have several common characteristics. For example, a shadow follows a highlight. A region of highlight in one beam indicates an object at that range. Then, the acoustic beam cannot travel further because the object blocks it. Therefore, a column of the sonar image has a specific pattern that the shadows follow the highlights.

The similarity between the sonar image and a template image can be calculated by sliding the template image size window and comparing each column of real image patch and template image, as below.

$$R_\theta = \max_{N\prime}\frac{1}{N_t}\sum_{n=1}^{N_t}\left(\max_{M\prime}\sum_{m=1}^{M_t} S(m+M', n+N')T(m,n)\right) \quad (6)$$

for $0 \leq M' \leq M - M_T$ and $0 \leq N' \leq N - N_T$, where the size of the sonar image is M by N, the size of the template image is M_T by N_T, and S and T denote the sonar image and template image, respectively.

Finally, the proposed method can detect the target object and recognize its angle by calculating the similarity for the template images of all angles and checking which angle has the highest similarity and whether the similarity value exceeds a certain threshold, as below.

$$\theta_{obj} = \underset{\theta}{\operatorname{argmax}} R_\theta \quad (7)$$

where $\max_{\theta} R_\theta$ exceeds a threshold value.

2.5 Application to the AUV's heading estimation

As one application of the proposed method, we present an estimation of the AUV's heading. The proposed method estimates the yaw angle of the underwater target object. The estimated angle is the relative bearing information of the object to the AUV. Therefore, if we assumed that the object is fixed, the AUV's heading angle can be estimated based on the object.

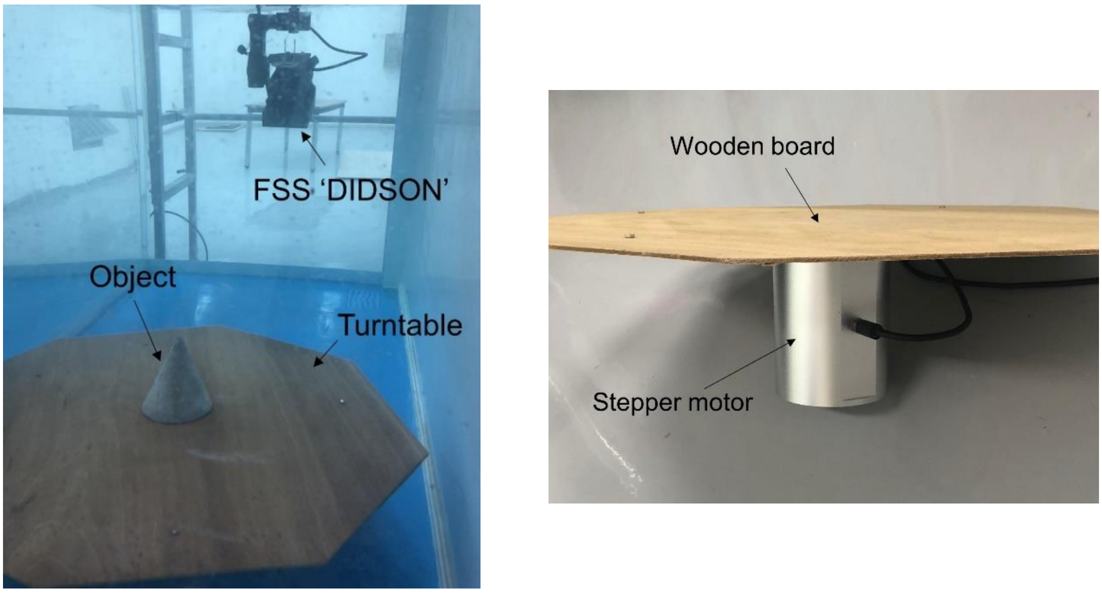
By installing landmarks in the field, the proposed method can be applied for heading angle correction of AUVs. When the AUV first identifies a landmark, the AUV estimates the landmark's angle and stores the information. If the AUV operates in the field and reencounters the same landmark, it can correct its yaw angle by identifying how much the landmark's yaw angle has changed. Likewise, the proposed method can be applied to other algorithms for AUV operation, such as bearing-based localization and mapping or pose estimation.

3. Experiments

3.1 Water tank experiments

3.1.1 Data acquisition

Water tank experiments were conducted to acquire a dataset to train the GAN for the style transfer and verify the proposed method, as shown in Fig. 7. The GAN training requires pairs of images simulated by the FSS simulator and real sonar images corresponding to their labels. A large number of and various types of images are needed to make the GAN produce high-quality output. Moreover, to verify the proposed yaw angle estimation method quantitatively, sonar images of an object according to the yaw angles are required.



(a) Experimental setup (b) Design of the turntable

Fig. 7 Water tank experiment for data acquisition

Table 1 Setup for the water tank experiment

Parameter	Value
Dimension (width x length x height)	1.35 m × 3 m × 1.7 m
FSS position from the center of the turntable (x, y, z)	1.6 m, 0 m, 0.9 m
Tilt angle of the FSS	15°, 20°, 25°
Object translation from the center of the turntable	0 m, 0.15 m

Table 2 Specifications of DIDSON

Features	Value
Field of view	0.42 ~ 46.25 m (in range) -14.5° ~ 14.5° (in azimuth) -7° ~ 7° (in elevation)
Operating frequency	1.8 MHz
Spreading angle	14°
Maximum resolution	0.3°
Image size	512 × 96
Depth rating	300 m

We designed a turntable to efficiently acquires various types of images and images according to the yaw angles of the object instead of changing the viewing angle of the sonar sensor. The turntable can rotate an object placed on it to the desired angle using a waterproof stepping motor. Moreover, the object and the background in images could be separated well by assembling a wooden board on the motor.

Objects of various shapes and materials, such as a plastic ball, concrete cone, rubber tire, and clay bricks, were placed in multiple positions on the turntable, and the sonar images of those objects were taken while rotating the turntable. Dual-frequency identification sonar (DIDSON) was used for the FSS (Belcher *et al*. 2002). Tables 1 and 2 show the setup for the water tank experiment and specification of DIDSON, respectively.

We trained GAN using supervised learning. For the supervised learning of the GAN to transfer the style of roughly simulated images to realistic images, the images captured in the tank are used as labels, and corresponding simulated images are required as inputs. We synthesized input training images under the same condition of the tank experiment using the FSS simulator. The parameters of the FSS simulator such as r_{min}, r_{max}, θ_{min} and θ_{max} was set according to the specification of DIDSON. The number of sample rays K, which determines the simulated image's precision, was set to 1,000. We then modeled the three-dimensional (3D) shapes of the turntable and the objects used in the tank experiment with the same dimension as the object using computer-aided design (CAD). Finally, an image having the same semantic information as the real sonar image was simulated by placing the 3D shapes and the virtual sonar sensor in the simulator world.

Data augmentation techniques such as scaling and horizontal and vertical flipping were applied to the dataset obtained from the experiments and simulations. As a result, 2,204 training image pairs were acquired, as shown in Fig. 8.

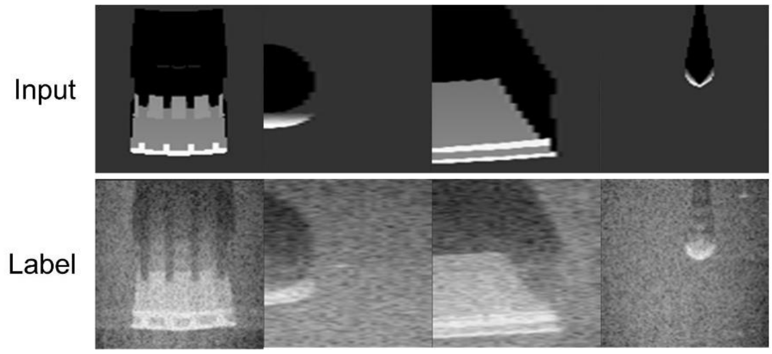


Fig. 8 Image pairs to train and validate GAN for the sytle transfer

(a) Type 1 (b) Type 2

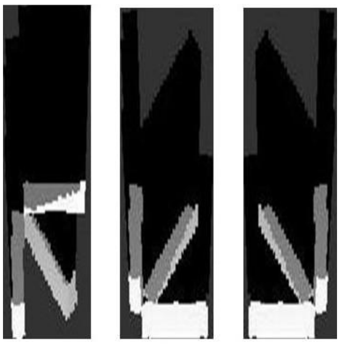
Fig. 9 Two types of objects to verify the proposed yaw angle estimation method

The test images for verifying the proposed yaw angle estimation were also taken through the water tank experiment. The bricks are stacked in two types to check whether the proposed method can distinguish only the target object from other objects, as shown in Fig. 9. Then, to quantitatively evaluate the angle estimation, the FSS images were taken while rotating the bricks by 10 degrees interval using the turntable.

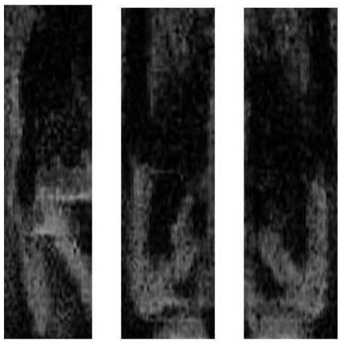
3.1.2 Yaw angle estimation result

We evaluated the accuracy of the estimated yaw angle of the underwater target object through the proposed pipeline. First, the target object's template images by the yaw angles are synthesized. The FSS simulator calculates the shape of the target object by yaw angles, as shown in Fig. 10(a). The GAN then synthesized realistic template images by applying the style transfer to the simulated shapes, as shown in Fig. 10(b). The number of template images or angle intervals can be determined according to the desired angle estimation resolution. In this paper, we synthesized 36 template images in 10-degree intervals.

The underwater target object's yaw angle can be estimated by calculating the similarity between sonar images and template images. Table 3 and Fig. 11 show the calculated similarity for all combinations between 36 sonar images taken at 10-degree intervals and 36 template image simulated at 10-degree intervals.



(a) Simulated shapes using the FSS simulator



(b) Realistic template images by applying the style transfer

Fig. 10 Samples of generated template images for yaw angle estimation

Table 3 Sampled yaw angle estimation result by matching with the generated template images

		Template Images of Type 1			
		50°	90°	200°	340°
Sonar images of Type 1	50°	0.4950	0	0	0
	90°	0	0.7267	0	0
	200°	0	0	0.5360	0.4952
	340°	0	0	0	0.8389
Sonar images of Type 2	50°	0.6827	0.4240	0.8777	0.8067
	90°	0.2410	0.1027	0.0338	0.3657
	200°	0	0	0.0837	0.2102
	340°	1	0.8783	0.9215	1

The similarity values calculated between sonar images captured object type 1 and template images of type 1 show a correlation when the value is high when the angle of the object and the template image coincide. As shown in Fig. 11, the proposed method could estimate the object's angle within 10 degrees by finding which angle of the template image is most similar to a given image. To verify whether the proposed method can distinguish the target object from other objects, we calculated the similarity between sonar images of type 2 and template images of type 1. The calculated similarity value was low overall, or there was no correlation between the angles of the template and the real object. Therefore, the proposed method can identify that the given object is not the target object by applying a threshold value and checking the distribution of the similarity values across the template images.

We also measured the similarity between sonar images and simulated images which style transfer using GAN was not applied to check its effect, as shown in Table 4. When the style transfer is not applied, there is a considerable difference between the template image and real sonar images. The difference makes the correlation between the angles of the object and the template image noisy. The proposed NST method could improve the yaw angle estimation precision by reducing the difference between the actual sonar images and the template images.

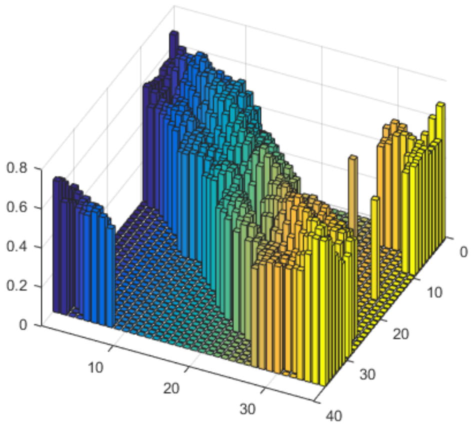


Fig. 11 Yaw angle estimation result

Table 4 Template matching results with the simulated images without applying the style transfer

		Template Images of Type 1 Before Applying Style Transfer			
		50°	90°	200°	340°
Sonar images of Type 1	50°	0.4968	0.3334	0.1811	0.2238
	90°	0.1605	0.5000	0.0670	0.1127
	200°	0.0971	0.0797	0.4372	0.1146
	340°	0.3425	0.1534	0.1944	0.5000

3.2 Field experiments

As an application of the proposed method, we tested the proposed method for heading estimation of an AUV in the field. We first designed a landmark, as shown in Fig. 12. Bricks were fixed on the wooden board in a 'C' shape. A board was then placed on the basket to distinguish the landmark from seaweeds or surrounding terrains easily. We installed the landmark on the seabed of Janggil-bay, Pohang, Republic of Korea.

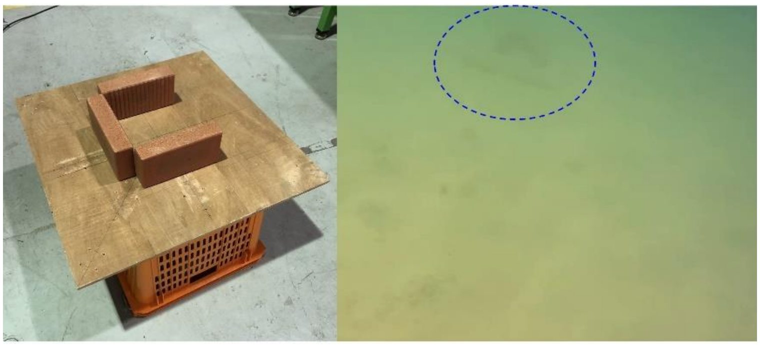


Fig. 12 Designed landmark for the bearing estimation of the AUV in the sea trials

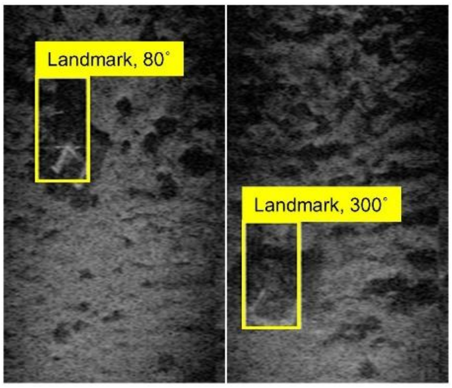


Fig. 13 Shape of landmark in sonar images captured in the sea trials

Then, the AUV equipped with FSS was operated on the field. The AUV captured sonar images of the seabed while moving. When the landmark was captured in the sonar sensor, the AUV could estimate bearing information from the landmark using the proposed method. As shown in Fig. 13, the proposed method could detect the landmark appearing at various positions in the sonar images and estimate the AUV's heading angle based on the detected landmark.

4. Conclusions

This paper proposed a method to estimate the yaw angle of underwater target objects. Due to the sonar imaging mechanism, the shape of an object changes drastically according to viewing angles. Taking advantage of this characteristic, we proposed a method to estimate an object's yaw angle by predicting the shape according to the angle in advance. The proposed method simulates the shape according to angles roughly using a sonar simulator. Then, GAN synthesized realistic template images by transferring styles of the simulated shapes. Finally, the object's yaw angle was identified by calculating similarities between the shape in real sonar image and each template image according to angle.

The estimation accuracy of the proposed method and the effect of style transfer using GAN were verified through water tank experiments. We also presented that the proposed method could be used for AUV operation in the field. Likewise, the proposed method can be applied to other AUV operations, such as object recognition using multi-view observing and bearing-based localization and mapping.

Acknowledgments

This research was a part of the project titled 'Gyeongbuk Sea Grant', funded by the Ministry of Oceans and Fisheries, Korea.

References

Belcher, E., Hanot, W. and Burch, J. (2002), "Dual-frequency identification sonar (DIDSON)", *Proceedings of the 2002 Interntional Symposium on Underwater Technology* (Cat. No. 02EX556).

Cho, H., Gu, J. and Yu, S.C. (2015), "Robust sonar-based underwater object recognition against angle-of-view variation", *IEEE Sensor. J.*, **16**(4), 1013-1025.

Etter, P.C. (1995), "Underwater acoustic modeling: principles, techniques and applications", CRC Press.

Hong, S. and Kim, J. (2020), "Three-dimensional Visual Mapping of Underwater Ship Hull Surface Using Piecewise-planar SLAM", *Int. J. Control Autom. Syst.*, **18**(3), 564-574.

Isola, P., Zhu, J.Y., Zhou, T. and Efros, A.A. (2017), "Image-to-image translation with conditional adversarial networks", *Proceedings of the IEEE conference on computer vision and pattern recognition*.

Johannsson, H., Kaess, M., Englot, B., Hover, F. and Leonard, J. (2010), "Imaging sonar-aided navigation for autonomous underwater harbor surveillance", *Proceedings of the 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*.

Karimanzira, D., Renkewitz, H., Shea, D. and Albiez, J. (2020), "Object Detection in Sonar Images", *Electronics*, **9**(7), 1180.

Kim, B. and Yu, S.C. (2017), "Imaging sonar based real-time underwater object detection utilizing adaboost method", *In 2017 IEEE Underwater Technology (UT)*, 1-5

Kim, J., Kim, T., Kim, J., Rho, S., Song, Y.W. and Yu, S.C. (2019), "Simulation and Feasibility Test of Mini-ROVs with AUV for the Manipulation Purpose", *In OCEANS 2019 MTS/IEEE SEATTLE*, 1-6.

Kim, J., Sung, M. and Yu, S.C. (2018a), "Development of simulator for autonomous underwater vehicles utilizing underwater acoustic and optical sensing emulators", *Proceedings of the 2018 18th International Conference on Control, Automation and Systems (ICCAS)*.

Kim, T., Kim, J. and Byun, S.W. (2018b), "A comparison of nonlinear filter algorithms for terrain-referenced underwater navigation", *Int. J. Control Autom. Syst.*, **16**(6), 2977-2989.

Lee, M., Kim, J. and Yu, S.C. (2019), "Robust 3D Shape Classification Method using Simulated Multi View Sonar Images and Convolutional Nueral Network", *In OCEANS 2019-Marseille*, 1-5.

Loc, M.B., Choi, H.S., Seo, J.M., Baek, S.H. and Kim, J.Y. (2014), "Development and control of a new AUV platform", *Int. J. Control, Autom. Syst.*, **12**(4), 886-894.

Maki, T., Horimoto, H., Ishihara, T. and Kofuji, K. (2020), "Tracking a Sea Turtle by an AUV with a Multibeam Imaging Sonar: Toward Robotic Observation of Marine Life", *Int. J. Control Autom. Syst.*, **18**(3), 597-604.

Myers, V. and Williams, D.P. (2011), "Adaptive multiview target classification in synthetic aperture sonar images using a partially observable Markov decision process", *IEEE J. Oceanic Eng.*, **37**(1), 45-55.

Palmese, M. and Trucco, A. (2006), "Acoustic imaging of underwater embedded objects: Signal simulation for three-dimensional sonar instrumentation", *IEEE T. Instrum. Measurement*, **55**(4), 1339-1347.

Ronneberger, O., Fischer, P. and Brox, T. (2015), "U-net: Convolutional networks for biomedical image segmentation", *Proceedings of the International Conference on Medical image computing and computer-assisted intervention*.

Tang, Z., Wang, Z., Lu, J., Ma, G. and Zhang, P. (2019), "Underwater robot detection system based on fish's lateral line", *Electronics*, **8**(5), 566.

Yu, S.C. (2008), "Development of real-time acoustic image recognition system using by autonomous marine vehicle", *Ocean Eng.*, **35**(1), 90-105.

MK